The background features a light beige color with several stylized, overlapping illustrations of newspapers and documents. These elements are rendered in a simple, hand-drawn style with brown outlines and some internal hatching to represent text columns. They are scattered around the central white text box, creating a layered, collage-like effect.

# **TruthNet: Developing a Natural Language Model for Fake News Detection**

---

By Sarah Roe

# Table of Contents

**01**

**Why?**

Why is this important?

**02**

**How? + What?**

How were models were created and what were the findings?

**03**

**Where?**

Where do we go from here?



01

**Why?**

---

**The ability to  
recognize factual  
news from  
misinformation is  
CRUCIAL!!**



# Problem Statement



Misinformation can have grave consequences



A classification model can aid in identifying real  
from fake



02

**How? + What?**

---

# Types of Errors? Which is worse?

**Positive:** The Onion

**Negative:** World News

Type I – False Positives

Type II – False Negatives

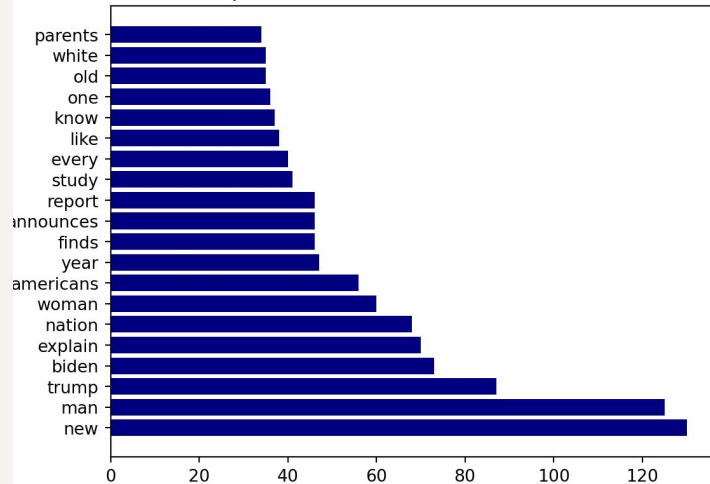
**Predicted Onion,  
Actually News**

**Predicted News,  
Actually Onion**

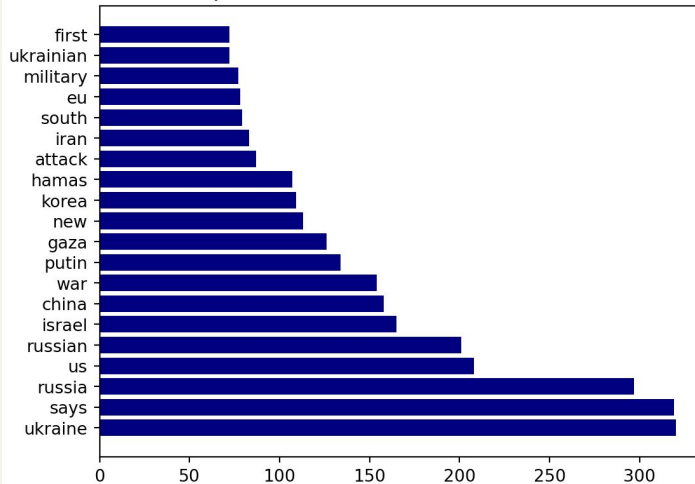
**Type II Error (False Negatives) are worse  
Want to optimize for Recall**

# Top 20 Words

Top 20 Words from The Onion Subreddit



Top 20 Words from World News Subreddit







# **Random Forest**

# Best Parameters + Metrics

## Tf-IDF:

- Stop Words: English + thread, worldnews, live, and onion
- Ngram range: (1, 1)

## Random Forest:

- Max Depth: None
- Max Features: 1

**Accuracy: 0.934**

**Recall: 0.92**

# Confusion Matrix

	<b>Pred World News</b>	<b>Pred The Onion</b>
<b>Actual World News</b>	569	33
<b>Actual Onion</b>	37	423

**Had more False Negatives and didn't perform  
much better with False Positives**


# What did it get wrong?



Biden Signs Executive Order to Deport All 340 Million Americans and Start From Scratch



Tiny Algae-made Robots Treat Inflammatory Bowel Disease without Drug



# **Logistic Regression + Sentiment Analysis**

# Best Parameters + Metrics

## Count Vectorizer

- Stop Words: English + thread, worldnews, live, and onion
- Ngram\_range: (1, 1)
- Minimum DF: 1

## Logistic Regression

- Penalty: l2
- C: 1.072

**Accuracy: 0.942**

**Recall: 0.94**

# Confusion Matrix

	<b>Pred World News</b>	<b>Pred The Onion</b>
<b>Actual World News</b>	569	33
<b>Actual Onion</b>	29	431

**Performed better with False Negatives and  
remained the same with False Positives**

# What did it get wrong?



IDF Soldier Recounts Harrowing, Heroic War Story of Killing 8-Month-Old Child



Spain Introduces Porn Passport to Stop Kids from Watching Smut



The background features a light beige color with several overlapping paper elements. A large, white, rounded rectangular sheet is the central focus. To its right, a portion of a newspaper is visible, showing columns of text represented by wavy lines and some solid black rectangular blocks. Below the main white sheet, another newspaper clipping is partially visible. The overall style is minimalist and illustrative.

03

**Where?**

---

# Implementation?

- Enhance media literacy
- Promote responsible consumption of information
- Continuous adaptation in order to conform to the evolving semantic and cultural nuances.
- Future modeling should involve using full articles for analysis

**Still had a 6% inaccuracy rate**



# Going Forward

We are at a moment in time where the divide between satire and reality have become increasingly blurred.

While relying on technology for almost everything in life is nice, it is so important that we not only continue to develop and improve our own critical thinking skills, but also urge others to do the same.

# References

[reddit.com/r/TheOnion](https://reddit.com/r/TheOnion)

[reddit.com/r/worldnews](https://reddit.com/r/worldnews)

General Assembly DSB Unit 3 Lessons

# Thanks!

---

Do you have any questions?



**CREDITS:** This presentation template was created by [Slidesgo](#), and includes icons by [Flaticon](#), and infographics & images by [Freepik](#)

Please keep this slide for attribution

